Mesterséges Intelligencia MI

Miért működik a minták alapján történő tanulás?
Mennyi tanítómintára van szükség?
VKH

Pataki Béla

BME I.E. 414, 463-26-79
pataki@mit.bme.hu,
http://www.mit.bme.hu/general/staff/pataki

Induktív tanulás

A tanítás folyamata:

Kiinduló (tanító) mintahalmaz

$\{(\mathbf{x}_n, d_n)\}$, $n=1,\ldots,N$

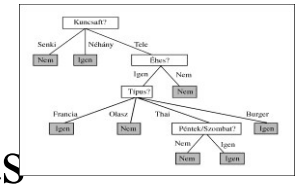
Például:

$\mathbf{x}_1$=[Kuncsaft=,Tele', Altern=,Igen', Bár=,Nem', ...]

d_1:VárniFog=,IGAZ'

h(x) hipotézis, mintákból leszűrendő általános szabály/tudás

például:



Tanítási algoritmus (hogyan építsük be a mintákban hordozott tudást az eszközbe)

Például: döntési fa kialakítása, növesztése

Induktív tanulás

A megtanított eszköz felhasználása

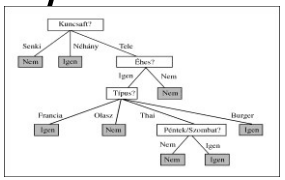
Új, ismeretlen szituáció leírása: $\mathbf{x}_{új}$

Például:

$\mathbf{x}_{új}$=[Kuncsaft=‚Senki', Altern=‚Nem', Bár=‚Nem', ...]

$h(\mathbf{x})$ hipotézis, mintákból leszűrt általános szabály/tudás

például:



Az új, ismeretlen szituációra ($\mathbf{x}_{új}$) javasolt válasz

Például:

$h(\mathbf{x}_{új})$= VárniFog=‚NEM'

Számítási tanulási elmélete

Hogyan és mennyire tudhatja valaki, hogy a tanulási algoritmusa olyan tudást eredményezett, amely helyesen fogja megjósolni a jövőben a válaszokat?

Az induktív tanulásnál:

honnan tudjuk, hogy a *h* (x) hipotézis jól közelíti az *f* (x) célfüggvényt, ha nem ismerjük *f*(x)-et?

Az *alapelv*

- bármely súlyosan hibás hipotézis már kis számú példa vizsgálata után is szinte biztosan „megbukik", mivel nagy valószínűséggel legalább egy helytelen eredményt fog jósolni, („találgatás")
- ezért **valószínűtlen, hogy súlyosan hibás lehet olyan hipotézis, amely egy kellően nagy tanuló példahalmazzal konzisztens** (a mintahalmaz összes elemére jó választ ad). **Természetesen 100% biztonság nincs, de nagyon valószínűtlen!**

Valószínűleg Közelítőleg Helyes (**Probably Approximately Correct**).
VKH-tanulás (**PAC**-learning)

Hány példára van szükség adott szintű hibához, és mennyire lehetünk biztosak abban, hogy teljesül a vállalt pontosság (hiba)?

X az összes lehetséges példák halmaza (lehet végtelen),

D a példák eloszlása.

H az összes lehetséges hipotézisek halmaza (komplexitás!)

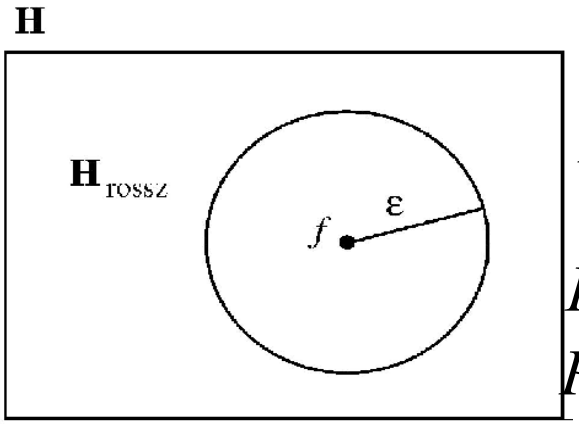
m a tanuló halmaz példáinak száma.

Legyen a keresett, valódi ***f*** függvény eleme **H**-nak. ***h*** hipotézis ***f*** függvényhez képesti **hibája** a ***D*** eloszlású példahalmazon: $e(h) = P(h(\mathbf{x}) \neq f(\mathbf{x}) \mid \mathbf{x} \in D)$

(*megjegyzés* 1: miért fontos odatennünk, hogy a *D* eloszlás feltételezésével?)
(*megjegyzés* 2: osztályozásnál $e(h)$ az osztályozási hibaarány, tehát a hipotézis – a tanított eszköz – más osztályba sorolja a mintát, mint ahova kéne)

h hipotézis **közelítőleg helyes**, ha $e(h) \leq \varepsilon$, ahol $\boldsymbol{\varepsilon}$ kicsi. Tehát legfeljebb ekkora az esély, hogy hibázik egy mintán, azaz a jó válasz esélye legalább:

Vizualizáció:



$$P(h_{VKH}(\mathbf{x}) \neq f(\mathbf{x}) \mid \mathbf{x} \in D) = e(h) < \varepsilon$$
$$P(h_{VKH}(\mathbf{x}) = f(\mathbf{x}) \mid \mathbf{x} \in D) \geq 1 - \varepsilon$$

$$P(h_{ROSSZ}(\mathbf{x}) \neq f(\mathbf{x}) \mid \mathbf{x} \in D) = e(h) > \varepsilon$$
$$P(h_{ROSSZ}(\mathbf{x}) = f(\mathbf{x}) \mid \mathbf{x} \in D) \leq 1 - \varepsilon$$

Mi annak a valószínűsége, hogy egy alapvetően rossz – nem közelítően helyes, nem VKH – hipotézis az összes (m db) mintával konzisztens?
Mivel h_{rossz} nem VKH, ezért annak valószínűsége, hogy bármelyik adott példára hibázik legalább $e(h_{rossz}) > \varepsilon$, tehát annak a valószínűsége, hogy egy adott mintára jó eredményt ad legfeljebb

$$P(h_{rossz}(\mathbf{x}) = f(\mathbf{x}) \mid \mathbf{x} \in D) \leq 1 - \varepsilon$$

Ez esetben annak valószínűsége, hogy <u>egyetlen</u> adott h_{rossz} mind az $\boldsymbol{m}$ példára jó választ ad: $P(h_{rossz}\ \textit{jó eredményt ad m példára}) \leq (1-\varepsilon)^m$

Annak valószínűsége, hogy az összesen $|\mathbf{H}_{rossz}|$ nem VKH hipotézis közül <u>valamelyik</u> konzisztens lesz mind az m mintánkkal:

$$P(h_{rossz} \in \mathbf{H}_{rossz} \text{konzisztens az } m \text{ mintánkkal}) \leq |\mathbf{H}_{rossz}|(1-\varepsilon)^m \leq |\mathbf{H}|(1-\varepsilon)^m$$

Hogyan korlátozhatjuk ezt a mintaszámmal?

$$P(h_{rossz} \in \mathbf{H}_{rossz} \text{konzisztens az } m \text{ mintánkkal}) \leq |\mathbf{H}|(1-\varepsilon)^m \leq \delta$$

$$|\mathbf{H}|(1-\varepsilon)^m \leq \delta$$

$(1-\varepsilon)^m \leq e^{-\varepsilon m}$ közelítés (Taylor-sor) $\Rightarrow$ $\boxed{\frac{1}{\varepsilon}\left(\ln\left(|\mathbf{H}|\right) - \ln\left(\delta\right)\right) \leq m}$

Ha egy tanuló algoritmus olyan hipotézist ad, amely ennyi megfelelően kiválasztott (véletlen mintavétel) példa esetén is konzisztens, akkor ennek a hipotézisnek legalább $1-\delta$ valószínűséggel a hibája legfeljebb ε. Más szavakkal **valószínűleg közelítőleg helyes (VKH)**.

A kívánt példaszám (ε , δés a hipotézistér mintakomplexitásának függvénye):

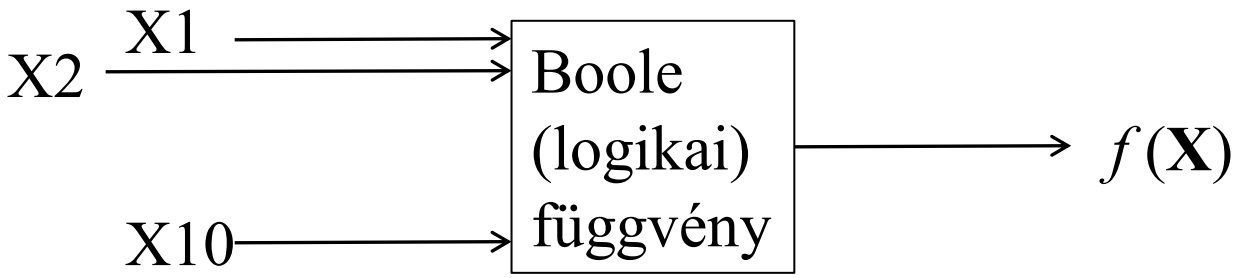
$$\frac{1}{\varepsilon}\left(\ln\left(\left|\mathbf{H}\right|\right)-\ln\left(\delta\right)\right)\leq m$$

Példa: ha $\left|\mathbf{H}\right|=10^7$ és 99% biztonsággal (1-δ=0,99, azaz δ=0,01) szeretnénk állítani, hogy a megtanított eszköz hibája nem lesz 5%-nál nagyobb, akkor

$$\frac{1}{0,05}\cdot\left(7\cdot\ln(10)-\ln(0,01)\right)=414,46<m$$

Tehát, ha m=415 véletlen minta mindegyikével konzisztens a tanított eszközünk, akkor ilyen biztonsággal garantálhatjuk a hibaszintet

$$Például: f(\mathbf{X}) = (X1 \wedge X2) \vee (X4 \wedge \neg X7) \wedge X5 \vee \neg X2 \ldots$$
$$\wedge X3 \vee (X10 \vee X3 \vee \neg X8)$$



Példa: Tanítsunk meg példák alapján egy 10-bemenetű Boole függvényt! Hány minta kell, ha max. 1%-os hibát 99,9% biztonsággal akarunk garantálni?

$$\left|\mathbf{H}\right| = 2^{2^{10}}$$

$$\frac{1}{0,01} \cdot \left(1024 \cdot \ln(2) - \ln(0,001)\right) = 71669 < m$$

Összesen hányféle bemeneti minta létezik?

Dilemma: ha nem korlátozzuk a tanuló algoritmus hipotéziseinek terét, akkor az algoritmus nem lesz képes tanulni (a szükséges tanítómintaszám $\rightarrow\infty$), ha viszont korlátozzuk, akkor lehet, hogy a valódi, keresett függvényt zárjuk ki.

Két kiút a csapdából:

- keresse az algoritmus a **legegyszerűbb** hipotézist, de legtöbb esetben a legegyszerűbb hipotézis számítása **kezelhetetlen**.
- legtöbb esetben nincs szükségünk pl. a Boole függvények teljes kifejező erejére, ennél kötöttebb – kevésbé komplex – eszközökkel is megoldhatók feladataink.